Journal of Cell Science 105, 799-806 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Protein localization to the nucleolus: a search for targeting domains in nucleolin

M. S. Schmidt-Zachmann* and E. A. Nigg†

Swiss Institute for Experimental Cancer Research (ISREC), Chemin des Boveresses 155, CH-1066 Epalinges, Switzerland

*Present address: German Cancer Research Center, Division of Cell Biology, Im Neuenheimer Feld 280, D-6900 Heidelberg, Germany
†Author for correspondence

SUMMARY

Nucleolin, a major nucleolar phosphoprotein, is presumed to function in rDNA transcription, rRNA packaging and ribosome assembly. Its primary sequence was highly conserved during evolution and suggests a multidomain structure. To identify structural elements required for nuclear uptake and nucleolar accumulation of nucleolin, we used site-directed mutagenesis to introduce point- and deletion-mutations into a chicken nucleolin cDNA. Following transient expression in mammalian cells, the intracellular distribution of the corresponding wild-type and mutant proteins was determined by indirect immunofluorescence microscopy. We found that nucleolin contains a functional nuclear localization signal (KRKKEMANKSAPEAKKKK) that conforms exactly to the consensus proposed recently for a bipartite signal (Robbins, J., Dilworth, S. M., Laskey, R. A. and Dingwall, C. (1991) *Cell* 64, 615-623). Concerning nucleolar localization, we found that the N-terminal 250 amino acids of nucleolin are dispensible, but deletion of either the centrally located RNA-binding motifs (the RNP domain) or the glycine/arginine-rich C terminus (the GR domain) resulted in an exclusively nucleoplasmic distribution. Although both of these latter domains were required for correct subcellular localization of nucleolin, they were not sufficient to target non-nucleolar proteins to the nucleolus. From these results we conclude that nucleolin does not contain a single, linear nucleolar targeting signal. Instead, we propose that the protein uses a bipartite NLS to enter the nucleus and then accumulates within the nucleolus by virtue of binding to other nucleolar components (probably rRNA) via its RNP and GR domains.

Key words: nucleolin, nuclear transport, NLS, nucleolar accumulation

INTRODUCTION

Protein import into the cell nucleus occurs through nuclear pore complexes (NPCs; Feldherr et al., 1984; for recent reviews see Ris, 1989; Akey, 1992; Jarnik and Aebi, 1991; Silver, 1991; Forbes, 1992; Dingwall and Laskey, 1992). These elaborate proteinaceous structures act not only as molecular sieves, allowing free diffusion of ions and small molecules, but also mediate the active transport of proteins and ribonucleoprotein particles. In order to enter the nucleus, proteins larger than about 60 kDa generally require a specific nuclear localization signal (NLS). NLSs have been identified for a number of viral and cellular proteins (for review see Garcia-Bustos et al., 1991; Dingwall and Laskey, 1991). They have been classified as either monopartite or bipartite, depending on whether or not stretches of basic residues are interrupted by a 10 amino acid spacer region (Robbins et al., 1991; reviewed by Dingwall and Laskey, 1991). The nuclear transport pathway can be separated into at least two steps, i.e. NLS-dependent targeting to the NPC, and ATP-dependent translocation through the NPC (Newmeyer and Forbes, 1988; Richardson et al., 1988). Recently, a number of NLS-binding proteins have been described (for review see Yamasaki and Lanford, 1992). These are proposed to function as cytoplasmic receptors for karyophilic proteins. They are implicated in delivering karyophilic proteins to the transport machinery of the NPC, but their exact roles remain to be determined. At present, little is known about the mechanisms that determine the localization of various nuclear proteins to specific intranuclear compartments. Although the sequence requirements for targeting lamins to the nuclear envelope are comparatively well understood (Loewinger and McKeon, 1988; Holtz et al., 1989; Krohne et al., 1989; Kitten and Nigg, 1991), and arginine/serine-rich domains have been implicated in localizing RNA-processing components to distinct subnuclear compartments (Li and Bingham, 1991), the mechanisms responsible for the association of proteins with other nuclear substructures remain largely unknown. One of the most conspicuous nuclear substructures is the nucleolus, the main site of ribosome biosynthesis in eukaryotic cells (reviewed by Hadjiolov, 1985; Scheer and Benavente, 1990). So called 'nucleolar targeting sequences', considered to be extended NLSs, have been

described for the heat-shock protein HSP70 (Munro and Pelham, 1984; Dang and Lee, 1989; Milarski and Morimoto, 1989) and for several viral proteins, including the TAT and Rev proteins of human immunodeficiency virus (HIV; Dang and Lee, 1989; Cochrane et al., 1990), and the Rex protein of human T-cell leukemia virus, type I (HTLV-I; Siomi et al., 1988). However, none of these proteins can be considered as a typical cellular component of the nucleolus, and the physiological role of the putative 'nucleolar targeting sequences' has not been clarified. The aim of the present study was to identify signals or domains required for nuclear import and nucleolar association of nucleolin (formerly termed C23). This protein is the major cellular constituent of nucleoli in exponentially growing cells (Bugler et al., 1982), and its abundance is correlated directly with the transcriptional activity of nucleoli (Escande-Géraud et al., 1985; Bouche et al., 1987). Nucleolin is a multifunctional protein involved in the organization of nucleolar chromatin (Olson and Thompson, 1983; Erard et al., 1988) and in the packaging of pre-rRNA (Herrera and Olson, 1986; Bugler et al., 1987). Moreover, the protein was shown to shuttle between the nucleus and the cytoplasm, suggesting a role in the transport of ribosomal proteins or preribosomal particles between the cytoplasm and the nucleolus (Borer et al., 1989). Nucleolin is a phosphoprotein (Olson et al., 1974). During interphase of the cell cycle, it is phosphorylated predominantly by casein kinase II (CKII) (Caizergues-Ferrer et al., 1987; Belenguer et al., 1989), whereas it is a substrate of the cell cycle-regulatory cdc2 kinase during mitosis (Peter et al., 1990; Belenguer et al., 1991). The primary sequence of nucleolin has been determined for several species (Lapeyre et al., 1987; Bourbon et al., 1988; Caizergues-Ferrer et al., 1989; Srivastava et al., 1989; Maridor and Nigg, 1990), and sequence comparisons reveal a high degree of evolutionary conservation: the protein consists of an N-terminal portion containing several acidic stretches; four RNA-binding motifs in the central region; and a glycine/arginine-rich domain at the very C terminus.

To determine the sequence requirements for nuclear import and nucleolar accumulation of nucleolin, we used a full length cDNA clone coding for the chicken protein to perform a detailed mutational analysis. The intracellular localization of wild type and mutant forms of nucleolin, as well as that of hybrids between parts of nucleolin and different reporter proteins, was then determined in a transient expression assay, using species-specific monoclonal antibodies for indirect immunofluroescence microscopy.

MATERIALS AND METHODS

Antibodies

The production and characterization of the chicken nucleolin-specific mAb I-8 was described earlier (Lehner et al., 1986; Borer et al., 1989). The mAb 9E10, used for detection of epitope-tagged protein constructs (Munro and Pelham, 1987) specifically recognizes a 10 amino acid peptide (EQKLISEEDL) derived from the human c-myc protein (Evan et al., 1985). All immunodetection experiments were carried out using either supernatants from hybridoma cultures (undiluted) or ascites fluids (diluted 1:1000). A guinea pig serum against the *Xenopus* nucleoplasmic protein N1/N2 (Kleinschmidt et al., 1985) was used at a dilution of 1:50.

Construction of mutant forms of chicken nucleolin

All constructs described below were derived from the wild-type chicken nucleolin cDNA cloned into the *Sma*I site of the pGEM-3Zf(−) vector (Promega), described by Maridor and Nigg (1990). For oligonucleotide-directed mutagenesis, the full-length nucleolin cDNA was subcloned into the double-stranded form of M13mp18. The corresponding single-stranded phage provided the template for second-strand synthesis using site-directed mutagenesis kits (Bio-Rad or Amersham) and appropiate mutant oligonucleotides as primers. All mutations were confirmed by sequencing, and cDNA inserts were re-cloned into pGEM plasmids.

The following mutant forms of chicken nucleolin were generated (see Fig. 2): in mutant M1, residues 256 to 260 (KRKK) were changed to QSNN, whereas in M2, residues 270 to 273 (KKKK) were changed to QQMN. In addition, a *Hin*dIII site was introduced into both M1 and M2 at nucleotide position 890; this allowed the construction of the double mutant M3. Two further mutants, M4 and M5 (not listed in Fig. 2), were constructed for convenience: in mutant M4, an internal *Sma*I site was introduced by changing nucleotides 930-933 from TGCT to CGGG, while in mutant M5 a *Pst*I site was introduced at nucleotides 1982-1985 (AAAG to TGCA). In mutant ΔGR, the codon for residue 631 was replaced by a stop-codon. For the deletion of the N-terminal part of nucleolin (ΔNt) a *Bam*HI site was introduced at nucleotide 835, and a new start-codon was created by changing lysine 251 to methionine. To create the mutants ΔRNP/GR and ΔRNP, respectively, M4 was cut with *Sma*I and *Hin*cII and either religated (to yield ΔRNP/GR), or blunt-end ligated to a fragment encoding the GR-domain (to yield ΔRNP); this latter fragment was obtained by digestion of M5 with *Pst*I.

Generation of epitope-tagged nucleolin constructs

Since some of the deletion mutants (ΔRNP; ΔRNP/GR) had lost the epitope for the mAb I-8 (which was mapped to a region close to RNP-domain II; M.S.Z. and E.A.N., unpublished results), they were tagged with an epitope derived from the human c-myc protein. A 100 bp-fragment (containing the myc-tag preceded by the 5′ untranslated region of human β globin) was excised by *Hin*dIII-*Eco*RI digestion from the plasmid pT7βTAG (Kobayashi et al., 1991) and cloned into a Bluescript expression vector (Stratagene). The resulting plasmid, in the following referred to as the BT-myc vector, contains several convenient restriction-sites downstream of the myc-tag and thus allows for in-frame insertion of appropriate cDNA fragments. To generate the tagged versions of wild-type nucleolin, ΔRNP and ΔRNP/GR, a *Hin*dIII-*Sma*I fragment derived from the BT-myc vector was blunt-end ligated into the *Ava*II site of the corresponding pGEM plasmids; as a consequence, 20 additional amino acids (SCSPRGSSAAAPAPPETAAI) were introduced between the myc-tag and these nucleolin sequences.

Construction of hybrid proteins

To generate a fusion protein between N1 and the C-terminal part of nucleolin (i.e. the RNP/GR domains), the full length cDNA coding for the N1 protein (Kleinschmidt et al., 1986) was subcloned into the *Eco*RI site of a Bluescript plasmid. Subsequently, a 1.5 kb-fragment derived from the nucleolin mutant M4 by digestion with *Sma*I and *Bam*HI was cloned into the *Nco*I site of this plasmid, resulting in the construct N1-RNP/GR. Pyruvate kinase (PK) containing the SV40-NLS at the 5′ end was isolated by *Xho*I-*Bam*HI digestion from the plasmid m30PKA (Kalderon et al., 1984), and blunt-end ligated into the *Eco*RI site of the pT7βTAG plasmid described above. The resulting myc-tagged version of NLS-PK was used further for the construction of the hybrid NLS-

PK-RNP/GR. For this purpose, NLS-PK was cut with *Asp*718, and a fragment coding for the RNP/GR-domains of nucleolin (derived from mutant M4 by *Sma*I-*Hin*dIII digestion) was inserted.

Transfection experiments

For transient expression in HeLa cells, the various cDNAs described above were subcloned into the *Hpa*I site of the mammalian expression vector pCMVneo (Krek and Nigg, 1991). Transfections, using 5 μg of DNA per 3.5 cm tissue culture dish, were carried out as described by Krek and Nigg (1991), using the method of Chen and Okayama (1987). For each transfection, time zero is defined as the moment when the DNA-Ca^{2+} precipitate was removed from the cells.

Indirect immunofluorescence microscopy

Transfected cells (grown on coverslips) were fixed for 7 min with 3% paraformaldehyde, 2% sucrose in phosphate-buffered saline (PBS: 137 mM NaCl; 2.7 mM KCl; 8.1 mM Na_2HPO_4; 1.5 mM KH_2PO_4, pH 7.2), and then processed as described by Krek and Nigg (1991). Incubations with primary and secondary antibodies were carried out for 15 min at room temperature. Secondary reagents were affinity-purified rhodamine-conjugated goat anti-mouse IgG (Cappel) and Texas Red-conjugated goat anti-guinea pig IgG (Dianova), respectively. Coverslips were mounted in 90% glycerol/10% 1 M Tris-HCl (pH 9.0), and cells were viewed with a Polyvar fluorescence microscope (Reichert-Jung), using ×40 or ×63 oil immersion objectives.

Sedimentation analysis, gel electrophoresis and immunoblotting

For sucrose gradient analysis, nucleolin was isolated from chicken hepatoma cells (DU249). These were cultured to confluency, as reported previously (Nakagawa et al., 1989). Cells were collected from a 10 cm Petri dish, washed once in PBS, and then homogenized by 30 strokes with a tight-fitting Dounce homogenizer in 1 ml PBS supplemented with 300 mM KCl, 1% aprotinin, 1 mM phenylmethylsulfonyl fluoride (PMSF). Following a 10 min incubation on ice, cellular debris was removed by centrifugation in a table-top minifuge (5 min, 14,000 r.p.m.). Of the resulting supernatant, 0.5 ml was layered on top of a 5% to 30% (w/v) linear sucrose gradient prepared in the above homogenization buffer. Reference proteins (bovine serum albumin, BSA (4.3 S); catalase (11.3 S); and thyroglobulin (16.5 S)) were applied to parallel gradients. These were then centrifuged at 36,000 r.p.m. in a Kontron TST 41.14 rotor for 16 h at 4°C. Fractions of 0.4 ml were collected, and proteins were precipitated with 20% trichloroacetic acid (final concentration). Following repeated washing with acetone, proteins were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting (Krek et al., 1992), using alkaline phosphatase-coupled anti-mouse Ig (Promega) as secondary antibodies.

RESULTS

Sedimentation behaviour of nucleolin

Before studying the subcellular localization of nucleolin mutants, it was important to determine whether or not nucleolin displays a propensity to oligomerize. The formation of heterotypic complexes between (endogenous) wild-type and (transfected) mutant proteins represents in fact a notorious problem with mutational analyses of subcellular protein trafficking (see, for instance, Loewinger and McKeon, 1988; Peculis and Gall, 1992). When nucleolin was isolated from cultured chicken cells and analyzed on linear sucrose gradients (Fig. 1), the bulk of the protein sedimented in fractions 4 to 6; for comparison, BSA (4.3 S) peaked in fraction 5, catalase (11. 3 S) in fraction 11, and thyroglobulin (16.5 S) in faction 18. No nucleolin was detectable by immunoblotting in later fractions (fractions 10-28; data not shown), indicating that the bulk of this 92 kDa protein exists predominantly as a monomer. We have also analyzed the sedimentation behaviour of [^{35}S]methionine-labeled nucleolin, following its in vitro translation in a rabbit reticulocyte lysate. Again, nucleolin displayed a sedimentation coefficient of 5 S, indicating that it does not readily form oligomeric assemblies (data not shown). These properties of nucleolin suggested that self-oligomerization would not seriously complicate the interpretation of localization studies. They set the stage for a systematic mutational analysis of the sequence requirements for nuclear and nucleolar accumulation of this protein.

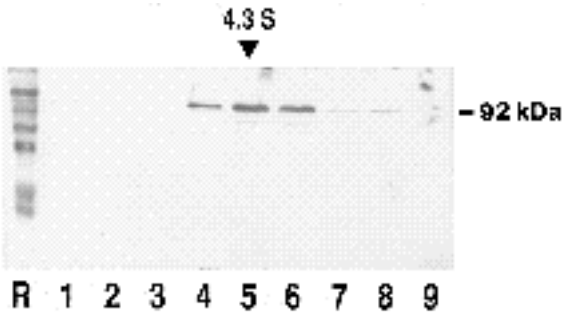

Fig. 1. Sedimentation behaviour of nucleolin. Nucleolin was extracted from DU249 cells and subjected to sucrose gradient centrifugation. Fractions were analyzed by SDS-PAGE and immunoblotting with the anti-nucleolin mAb I-8. Fraction numbers are indicated below the lanes (top of the gradient on the left). Reference proteins (lane R) were, from top to bottom: α2-macroglobuline (191 kDa), β-galactosidase (117 kDa), fructose-6-phosphate kinase (91.8 kDa), pyruvate kinase (72.7 kDa), fumarase (57.8 kDa), lactic dehydrogenase (40.8 kDa), triosephosphate isomerase (34.1 kDa). The arrowhead indicates the sedimentation of BSA, used as an S-value marker protein.

Identification of a bipartite NLS in nucleolin

The primary sequence of wild-type chicken nucleolin is shown schematically in Fig. 2. The major characteristics of this protein are four large acidic clusters within the N-terminus (A1 to A4), four RNA-binding motifs (RNP I to IV) in the central region, and a stretch rich in glycine and arginine residues (GR-domain) at the C terminus. Since inspection of the sequence suggested that a NLS might be located upstream of the four RNP-domains (Fig. 2, hatched area), the two basic clusters present in this region were mutated (as indicated in Fig. 2), either individually (M1; NLS_l, and M2; NLS_r) or in combination (M3; NLS_{l+r}). Wild type and putative NLS-mutants of nucleolin were then introduced into mammalian cells, and their intracellular distributions were monitored 24 h after transfection. When analyzed by indirect immunofluorescence microscopy with the chicken-specific mAb I-8, cells expressing wild-type nucleolin showed a bright nucleolar staining (Fig. 3a), demonstrating proper localization of the chicken protein in a heterologous

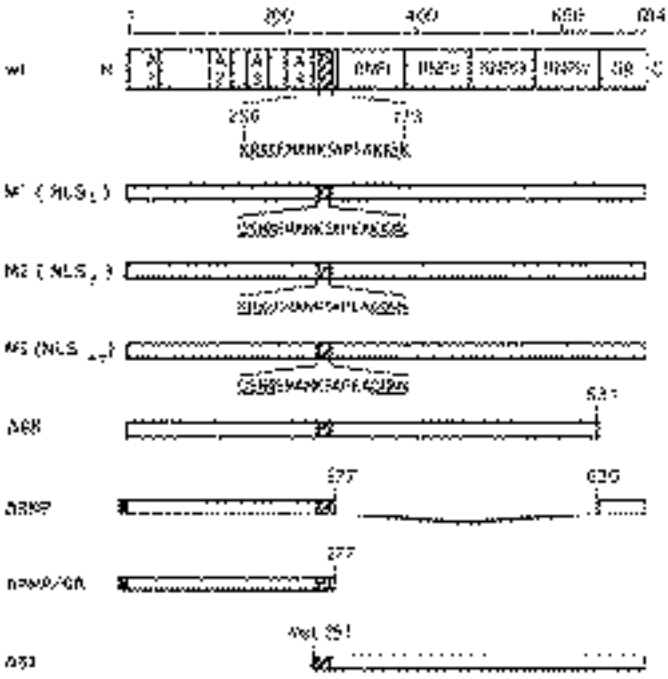

Fig. 2. Schematic summary of nucleolin mutants. For reference, structural organization of wild-type (wt) nucleolin is drawn to scale; indicated are the N-terminal acidic clusters (A1 to A4), the central RNA-binding motifs (RNP I-IV) and the C-terminal glycine/arginine-rich domain (GR). The hatched box denotes the position of the bipartite NLS identified in this study. Its protein sequence is shown enlarged, using the one-letter code; two basic clusters (left and right) are underlined. For all deletion mutants, the precise boundaries are indicated (as amino acid positions). In ΔRNP and ΔRNP/GR, the black bar denotes the myc-tag, while the internal deletion of ΔRNP is marked as a single line.

environment. In contrast, none of the putative NLS-mutants M1, M2 or M3, was able to accumulate in the nuclei of transfected cells; instead, the corresponding proteins remained almost exclusively cytoplasmic (Fig. 3b-d). Only a faint nucleolar staining was occasionally visible, possibly reflecting a very low level of piggy-back transport of mutant nucleolin with endogenous wild-type protein.

The above results were confirmed by injecting [^{35}S]methionine-labeled in vitro synthesized nucleolin into the cytoplasm of *Xenopus laevis* oocytes. After different incubation times, the oocytes were dissected manually, and the resulting nuclear and cytoplasmic fractions were analyzed by SDS-PAGE and autoradiography (data not shown). Whereas wild-type nucleolin was able to accumulate in the nucleus (to 50% after 16 h), none of the NLS-mutants M1, M2 or M3 was detectable in the nuclear fraction even after overnight incubation. From these results we conclude that nucleolin contains a typical bipartite nuclear location signal, as described originally for the histone-binding protein nucleoplasmin of *Xenopus laevis* (Robbins et al., 1991).

Accumulation in the nucleolus requires two structural elements of nucleolin

Having identified the NLS of nucleolin, we next investigated the possible subnuclear targeting functions of the different structural domains present in the protein. For this purpose, several deletion mutants were constructed (summarized in Fig. 2), and following their expression in HeLa cells, their subcellular distributions were determined (Fig. 4). While the presence of the myc-tag did not alter the distribution of wild-type nucleolin (data not shown), a deletion of the glycine/arginine stretch (ΔGR) prevented the protein from accumulating in nucleoli, and instead resulted in a uniformly nucleoplasmic distribution (Fig. 4a). The same localization was observed also for the mutant that had a complete GR-domain but lacked the RNP-domain (ΔRNP; Fig. 4b), and for the double-mutant lacking both RNP and GR-domains (ΔRNP/GR; data not shown). In contrast, the mutant that lacked the entire N terminus, including the four acidic domains and all phosphorylation sites identified so far, still localized efficiently to nucleoli (Fig. 4c).

Nucleolin does not contain a transferable nucleolar targeting signal

As shown above, nucleolar accumulation of nucleolin

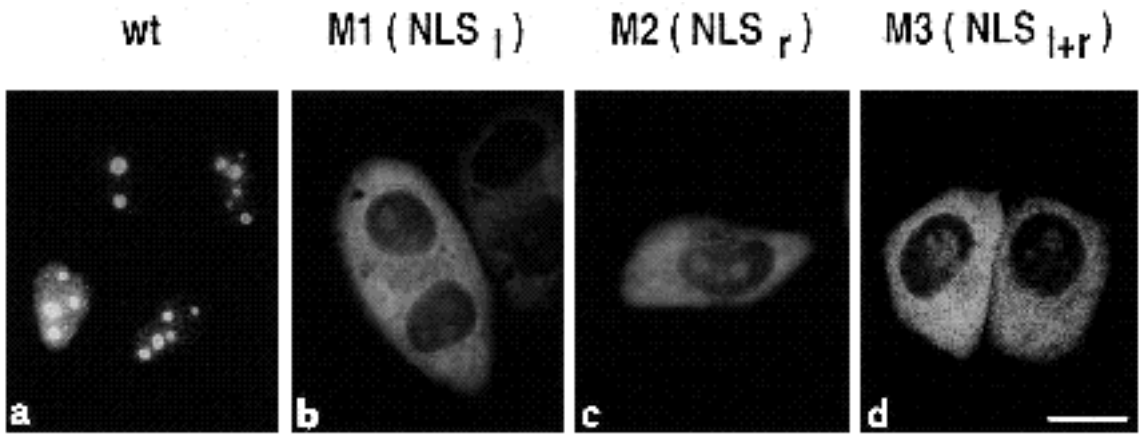

Fig. 3. Subcellular localization of wild-type and NLS-mutant nucleolin. At 24 h after transfection, HeLa cells were fixed and stained with the chicken-specific anti-nucleolin mAb I-8, followed by rhodamin-conjugated goat anti-mouse IgG antibodies.(a) wild-type nucleolin; (b-d) NLS-mutants; these are designated as described in Fig. 2. Bar in d, 20 μm.

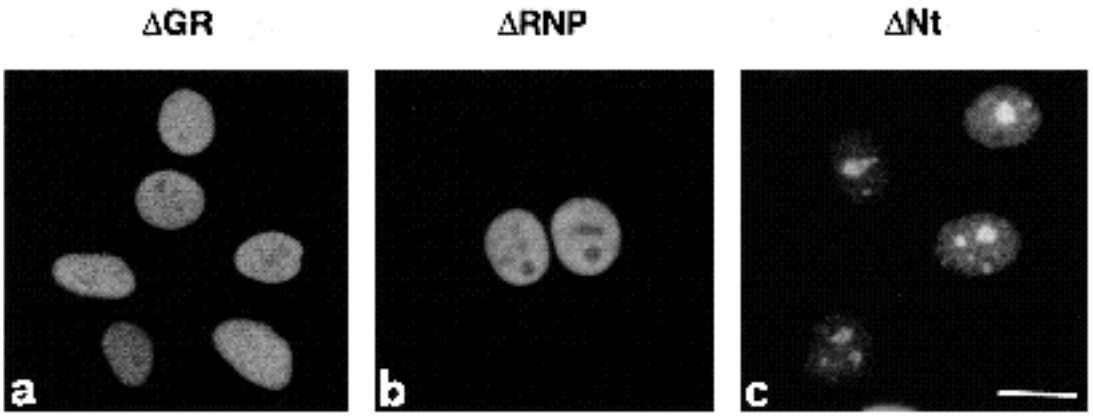

Fig. 4. Subcellular localization of nucleolin mutants lacking different structural domains. Transfection experiments and immunofluorescent staining were carried out as described in the legend to Fig. 3. Transfected cells were stained either with mAb I-8 (a,c) or with mAb 9E10 (b). Mutants are designated as described in Fig. 2. Note that removal of the very C terminus (ΔGR; a) or the internal deletion of the RNA-binding motifs (ΔRNP; b) resulted in an exclusively nucleoplasmic distribution of the mutant proteins. In contrast, the mutant missing the entire N terminus (ΔNt; c) was still able to localize to nucleoli. Bar in c, 15 μm.

requires the RNP as well as the GR-domain. To determine whether a combination of these domains would be sufficient to target non-nucleolar proteins to the nucleolus, two hybrid proteins were constructed (summarized in Fig. 5A). As a first reporter protein we used N1, a well-characterized histone-binding protein from *Xenopus laevis* (Kleinschmidt et al., 1986). Wild-type N1 contains a bipartite NLS and localizes to the nucleoplasm (Kleinschmidt and Seiter, 1988; see also Robbins et al., 1991). As a second reporter protein, we chose a completely artificial 'nuclear protein', namely a chicken pyruvate kinase (PK) fused to the NLS of SV40 T-antigen. This was done to minimize the possibility that the reporter protein itself would display strong affinities for nucleoplasmic binding sites. When analyzed by transfection, wild-type N1 (Fig. 5B, panel a) as well as NLS-PK (Fig. 5B, panel c) were present in the nucleoplasm, as expected. Fusion of the RNP/GR-domain to these proteins did not confer nucleolar localization to either the N1 protein (Fig. 5B, panel b) or the NLS-PK (Fig. 5B, panel d). These results indicate that the RNP/GR-domains are essential for the nucleolar accumulation of nucleolin, but are not sufficient to redirect hybrid-proteins to the nucleolus. Similar conclusions have been reached independently by Messmer and Dreyer (1993).

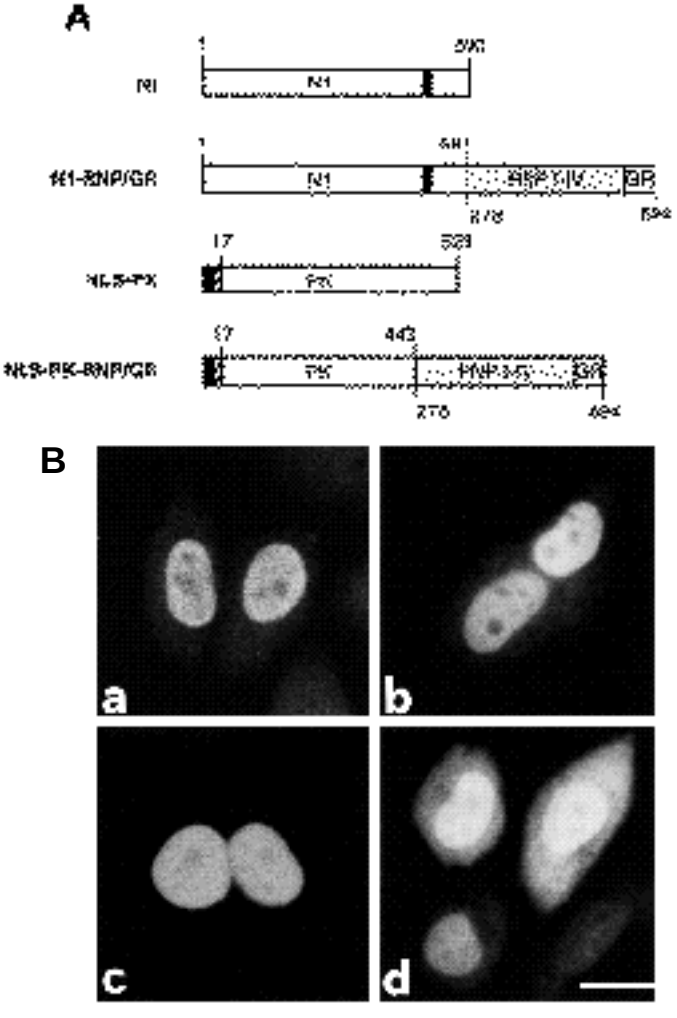

Fig. 5. Construction and analysis of hybrid proteins:
(A) schematic description of fusion protein constructs; amino acid positions at termini and junctions are indicated by numbers. The darkly shaded box marks the position of the bipartite NLS of *Xenopus laevis* protein N1 (Kleinschmidt and Seiter, 1988). The cytoplasmic protein pyruvate kinase (PK), containing the SV40-NLS (hatched area), was myc-tagged (indicated by the black box). Both reporter proteins were fused to the RNP/GR domains of nucleolin (shown as lightly shaded boxes). (B) Subcellular localization of hybrid proteins after expression in HeLa cells. Transfected cells were stained after 24 h, either with antiserum against N1 (a,b) or with anti-myc tag mAb 9E10 (c,d). Notice that both wild-type N1 protein (a) and the hybrid protein N1-RNP/GR (b) display an undistinguishable nucleoplasmic distribution. Likewise, NLS-PK (c) as well as the corresponding hybrid protein NLS-PK-RNP/GR (d), were nucleoplasmic, with no evidence for nucleolar accumulation. Although the RNP/GR-domain seemed to slow down the nuclear uptake of NLS-PK, as indicated by an increased cytoplasmic staining, complete nuclear uptake was seen by 48 h after transfection (data not shown). Bar in d, 15 μm.

DISCUSSION

While the structural signals responsible for nuclear uptake of proteins are comparatively well characterized, it remains to be determined to what extent distinct amino acid sequence motifs govern the targeting of proteins to precise subnuclear compartments. To address this issue, we have analyzed the subnuclear localization of various mutants of the major nucleolar protein nucleolin. We demonstrate that the nuclear uptake of nucleolin is mediated by a NLS of the bipartite type. Its nucleolar accumulation, however, is not controlled by a 'signal' sequence, as was claimed previously for viral proteins (e.g. Siomi et al., 1988). Instead, efficient localization of nucleolin to the nucleolus depends on the presence of both RNA-binding domains and a glycin/arginine-rich C terminus. The bipartite NLS of chicken nucleolin was mapped to residues 256-273, just upsteam of the RNP-domain. Its sequence KRKKEMANKSAPEAKKKK conforms very well to the consensus proposed for bipartite NLSs (Robbins et al., 1991; Dingwall and Laskey, 1991), and we have demonstrated that both basic domains of this bipartite signal (underlined) are required for nuclear uptake of nucleolin. Of the several structural domains present in nucleolin, only the N-terminus was found to be dispensible for nucleolar accumulation. This N-terminus contains about 250 amino acids; its precise function remains to be determined, but it has been proposed to confer on nucleolin a high affinity for histone H1, and to serve in the displacement of histone H1 from nucleolar chromatin (Erard et al., 1988; Erard et al., 1990). In contrast, both the RNP motifs and the glycine/arginine-rich C terminus were shown here to be required for the nucleolar accumulation of nucleolin. The four RNP motifs were previously implicated in mediating the binding of nucleolin to the 5′ external transcribed spacer of ribosomal RNA (Bugler et al., 1987, Ghisolfi et al., 1990). The C terminal GR-domain, approximately 70 amino acids long, is rich in glycines, with interspersed dimethylarginine and phenylalanine residues (Lapeyre et al., 1986). Its exact role in vivo is presently unknown, but it is interesting that a combination of RNP and GR-domains is not exclusive to nucleolin. Such domains are found also in the nucleolar proteins fibrillarin (Lapeyre et al., 1990), GAR1 (Girard et al., 1992) and NSR 1 (Lee et al., 1991), as well as in the non-nucleolar hnRNP protein A1 (Cobianchi et al., 1986; Burd et al., 1989). Recent in vitro data indicate that GR-domains may bind non-specifically to RNAs, thereby unfolding them to allow efficient and specific binding of the RNP-domains (Ghisolfi et al., 1992a,b). Sequences responsible for nucleolar accumulation have previously been studied in other proteins. In the case of the stress-protein HSP70, somewhat conflicting results have been reported: in *Drosophila* HSP70, a N-terminal sequence of 18 amino acids was described to be required for nucleolar localization (Munro and Pelham, 1984; Dang and Lee, 1989), but in the human homolog, an essential region was mapped to the C-terminal half of the protein (Milarski and Morimoto, 1989). A C-terminal domain of about 24 amino acids was also implicated in the nucleolar localization of the nucleolar protein NO38 (Peculis and Gall, 1992), and a C-terminal acidic domain as well as a DNA-binding region were found to be necessary for the nucleolar accumulation of the transcription factor UBF (Maeda et al., 1992). Finally, short 'nucleolar targeting signals' have been described for several viral proteins, including Rex, Rev and Tat (for references see Introduction). The domains shown here to be required for the nucleolar accumulation of nucleolin cover about two thirds of the entire protein. Also, we emphasize that transfer of the RNP- and GR-domains of nucleolin to reporter proteins did not result in targeting of the resulting hybrid proteins to the nucleolus, suggesting that nucleolar localization may require appropriate folding of rather extensive protein domains (see also Messmer and Dreyer, 1993). These latter results contrast with the finding that the relatively short 'nucleolar targeting signals' of certain viral proteins could confer nucleolar localization to β-galactosidase (reviewed by Hatanaka, 1991). However, the difference between these results may be more apparent than real. Recent studies in fact indicate that the 'nucleolar targeting signal' of the viral Tat protein binds to an RNA-stem-loop structure in the HIV long terminal repeat (Weeks et al., 1990; Cordingly et al., 1990). If similar binding sites were present in nucleolar rRNA, this could account for the nucleolar accumulation of the Tat protein. Hence, as proposed here for the RNP/GR-domains of nucleolin, the nucleolar accumulation of viral Tat protein may not be 'signal-mediated' but depend on RNA binding.

In summary, our studies lead us to conclude that there is no consensus signal sequence for targeting proteins to the nucleolus. Instead, we propose that accumulation of proteins in the nucleolus results from specific binding interactions between these proteins and other nucleolar components, particularly rDNA, rRNA, and possibly protein constituents of a nucleolar matrix structure.

We thank Drs J. Kleinschmidt and W. Richardson for kind gifts of N1 and pyruvate-kinase plasmids, respectively. We also thank Drs C. Dargemont and H. Hennekes for helpful and stimulating discussions. This work was supported by an EMBO fellowship (to Marion S. Schmidt-Zachmann) and grants from the Swiss National Science Foundation (31-33615.92) and the Swiss Cancer League (FOR 205) to Erich A. Nigg.

REFERENCES

Akey, C. W. (1992). The nuclear pore complex: a macromolecular transport assembly. In *Nuclear Trafficking* (ed. C. Feldherr), p. 370. New York: Academic Press.

Belenguer, P., Baldin, V., Mathieu, C., Prats, H., Bensaid, M., Bouche, G. and Amalric, F. (1989). Protein kinase NII and the regulation of rDNA transcription in mammalian cells. *Nucl. Acids Res.* **17**, 6625-6637.

Belenguer, P., Caizergues-Ferrer, M., Labbé, J.-C., Dorée, M. and Amalric, F. (1991). Mitosis-specific phosphorylation of nucleolin by p34cdc2 protein kinase. *Mol. Cell. Biol.* **10**, 3607-3618.

Borer, R. A., Lehner, C. F., Eppenberger, H. M. and Nigg, E. A. (1989). Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell* **56**, 379-390.

Bouche, G., Gas, N., Prats, H., Baldin, V., Tauber, J. P., Teissie, J. and Amalric, F. (1987). Basic fibroblast growth factor enters the nucleolus and stimulates the transcription of ribosomal genes in ABAE cells undergoing $G_0 \rightarrow G_1$ transition. *Proc. Nat. Acad. Sci. USA* **84**, 6770-6774.

Bourbon, H. M., Lapeyre, B. and Amalric, F. (1988). Structure of the mouse nucleolin gene: the complete sequence reveals that each RNA binding domain is encoded by two independent exons. *J. Mol. Biol.* **200**, 627-638.

Bugler, B., Bourbon, H., Lapeyre, B., Wallace, M. O., Chang, J.-H., Amalric, F. and Olson, M. O. J. (1987). RNA binding fragments from nucleolin contain the ribonucleoprotein consensus sequence. *J. Biol. Chem.* **262**, 10922-10925.

Bugler, B., Caizergues-Ferrer, M., Bouche, G., Bourbon, H. and Amalric, F. (1982). Detection and localization of a class of proteins immunologicaly related to a 100-kDa nucleolar protein. *Eur. J. Biochem.* **128**, 475-480.

Burd, C. G., Swanson, M. S., Görlach, M. and Dreyfuss, G. (1989). Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. *Proc. Nat. Acad. Sci. USA* **86**, 9788-9792.

Caizergues-Ferrer, M., Belenguer, P., Lapeyre, B., Amalric, F., Wallace, M. O. and Olson, M. O. (1987). Phosphorylation of nucleolin by a nucleolar type NII protein kinase. *Biochemistry* **26**, 7876-7883.

Caizergues-Ferrer, M., Mariottini, P., Curie, C., Lapeyre, B., Gas, N., Amalric, F. and Amaldi, F. (1989). Nucleolin from Xenopus laevis: cDNA cloning and expression during development. *Genes Dev.* **3**, 324-333.

Chen, C. and Okayama, H. (1987). High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745-2752.

Cobianchi, F., SenGupta, D. N., Zmudzka, B. Z. and Wilson, S. H. (1986). Structure of rodent helix-destabilizing protein revealed by cDNA cloning. *J. Biol. Chem.* **261**, 3536-3543.

Cochrane, A. W., Perkins, A. and Rosen, C. A. (1990). Identification of sequences important in the nucleolar localization of human immunodeficiency virus rev: relevance of nucleolar localization to function. *J. Virol.* **64**, 881-885.

Cordingley, M. G., LaFemina, R. L., Callahan, P. L., Condra, J. H., Sardana, V. V., Graham, D. J., Nguyen, T. M., LeGrow, K., Gotlib, L., Schlabach, A. J. and Colonno, R. J. (1990). Sequence-specific interaction of Tat protein and Tat peptides with the transactivation-responsive sequence element of human immunodeficiency virus type 1 in vitro. *Proc. Nat. Acad. Sci. USA* **87**, 8985-8989.

Dang, C. V. and Lee, W. M. F. (1989). Nuclear and nucleolar targeting sequences of c-erb-A, c-myb, N-myc, p53, HSP70, and HIV tat proteins. *J. Biol. Chem.* **264**, 18019-18023.

Dingwall, C. and Laskey, R. A. (1991). Nuclear targeting sequences - a consensus? *Trends Biochem. Sci.* **16**, 478-481.

Dingwall, C. and Laskey, R. A. (1992). The nuclear membrane. *Science* **258**, 942-947.

Erard, M. S., Belenguer, P., Caizergues-Ferrer, M., Pantaloni, A. and Amalric, F. (1988). A major nucleolar protein, nucleolin, induces chromatin decondensation by binding to histone H1. *Eur. J. Biochem.* **175**, 525-530.

Erard, M., Lakhdar-Ghazal, F. and Amalric, F. (1990). Repeat peptide motifs which contain beta-turns and modulate DNA condensation in chromatin. *Eur. J. Biochem.* **191**, 19-26.

Escande-Géraud, M. L., Azum, M. C., Tichadou, J. L. and Gas, N. (1985). Correlation between rDNA transcription and distribution of a 100 kDa nucleolar protein in CHO cell. *Exp. Cell Res.* **161**, 353-363.

Evan, G. I., Lewis, G. K., Ramsay, G. and Bishop, J. M. (1985). Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell. Biol.* **5**, 3610-3616.

Feldherr, C. M., Kallenbach, E. and Schultz, N. (1984). Movement of a karyophilic protein through the nuclear pores of oocytes. *J. Cell Biol.* **99**, 2216-2222.

Forbes, D. J. (1992). Structure and function of the nuclear pore complex. *Annu. Rev. Cell Biol.* **8**, 495-527.

Garcia-Bustos, J., Heitman, J. and Hall, M. N. (1991). Nuclear protein localization. *Biochim. Biophys. Acta* **1071**, 83-101.

Ghisolfi, L., Joseph, G., Amalric, F. and Erard, M. (1992a). The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-Helix-destabilizing properties. *J. Biol. Chem.* **267**, 2955-2959.

Ghisolfi, L., Joseph, G., Erard, M., Escoubas, J.-M., Mathieu, C. and Amalric, F. (1990). Nucleolin - pre-rRNA interactions and preribosome assembly. *Mol. Biol. Rep.* **14**, 113-114.

Ghisolfi, L., Kharrat, A., Joseph, G., Amalric, F. and Erard, M. (1992b). Concerted activities of the RNA recognition and the glycine-rich C-terminal domains of nucleolin are required for efficient complex formation with pre-ribosomal RNA. *Eur. J. Biochem.* **209**, 541-548.

Girard, J.-P., Lehtonen, H., Caizergues-Ferrer, M., Amalric, F., Tollervey, D. and Lapeyre, B. (1992). GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. *EMBO J.* **11**, 673-682.

Hadjiolov, A. A. (1985). The nucleolus and ribosome biogenesis. *Cell Biol. Mon.* **12**, 1-268.

Hatanaka, M. (1991). Nucleolar targeting signals (NOS) of HIV. *Genetic Structure and Regulation of HIV* (ed. W. A. Haseltine and F. Wong-Staal), pp. 264-287. Raven Press Ltd, New York.

Herrera, A. H. and Olson, M. O. (1986). Association of protein C23 with rapidly labeled nucleolar RNA. *Biochemistry* **25**, 6258-6264.

Holtz, D., Tanaka, R. A., Hartwig, J. and McKeon, F. (1989). The CaaX motif of lamin A functions in conjunction with the nuclear localization signal to target assembly to the nuclear envelope. *Cell* **59**, 969-977.

Jarnik, M. and Aebi, U. (1991). Toward a more complete 3-D structure of the nuclear pore complex. *J. Struct. Biol.* **107**, 291-308.

Kalderon, D., Roberts, B. L., Richardson, W. D. and Smith, A. E. (1984). A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499-509.

Kitten, G. T. and Nigg, E. A. (1991). The CaaX motif is required for isoprenylation, carboxyl methylation and nuclear membrane association of lamin B2. *J. Cell Biol.* **113**, 13-23.

Kleinschmidt, J. A., Dingwall, C., Maier, G. and Franke, W. W. (1986). Molecular characterization of a karyophilic, histone-binding protein: cDNA cloning, amino acid sequence and expression of nuclear protein N1/N2 of Xenopus laevis. *EMBO J.* **5**, 3547-3552.

Kleinschmidt, J. A., Fortkamp, E., Krohne, G., Zentgraf, H. and Franke, W. W. (1985). Co-existence of two different types of soluble histone complexes in nuclei of Xenopus laevis oocytes. *J. Biol. Chem.* **260**, 1166-1176.

Kleinschmidt, J. A. and Seiter, A. (1988). Identification of domains involved in nuclear uptake and histone binding of protein N1 of Xenopus laevis. *EMBO J.* **7**, 1605-1614.

Kobayashi, H., Golsteyn, R., Poon, R., Stewart, E., Gannon, J., Minshull, J., Smith, R. and Hunt, T. (1991). Cyclins and their partners during Xenopus oocyte maturation. *Cold Spring Harb. Symp. Quant. Biol.* **LVI**, 437-447.

Krek, W., Maridor, G. and Nigg, E. A. (1992). Casein kinase II is a predominantly nuclear enzyme. *J. Cell Biol.* **116**, 43-55.

Krek, W. and Nigg, E. A. (1991). Mutations of p34cdc2 phosphorylation sites induce premature mitotic events in HeLa cells: evidence for a double block to p34cdc2 kinase activation in vertebrates. *EMBO J.* **10**, 3331-3341.

Krohne, G., Waizenegger, I. and Höger, T. H. (1989). The conserved carboxy-terminal cysteine of nuclear lamins is essential for lamin association with the nuclear envelope. *J. Cell Biol.* **109**, 2003-2011.

Lapeyre, B., Amalric, F., Ghaffari, S. H., Venkatarama Rao, S. V., Dumbar, T. S. and Olson, M. O. J. (1986). Protein and cDNA sequence of a glycine-rich, dimethylarginine-containing region located near the carboxyl-terminal end of nucleolin (C23 and 100 kDa). *J. Biol. Chem.* **261**, 9167-9173.

Lapeyre, B., Bourbon, H. and Amalric, F. (1987). Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. *Proc. Nat. Acad. Sci. USA* **84**, 1472-1476.

Lapeyre, B., Mariottini, P., Mathieu, C., Ferrer, P., Amaldi, F., Amalric, F. and Caizergues-Ferrer, M. (1990). Molecular cloning of Xenopus fibrillarin, a conserved U3 small nuclear ribonucleoprotein recognized by antisera from humans with autoimmune disease. *Mol. Cell. Biol.* **10**, 430-434.

Lee, W. C., Xue, Z. and Mélèse, T. (1991). The NSR1 gene encodes a protein that specifically binds nuclear localization sequences and has two RNA recognition motifs. *J. Cell Biol.* **113**, 1-12.

Lehner, C. F., Eppenberger, H. M., Fakan, S. and Nigg, E. A. (1986). Nuclear substructure antigens: monoclonal antibodies against components of nuclear matrix preparations. *Exp. Cell Res.* **162**, 205-219.

Li, H. and Bingham, P. M. (1991). Arginine/serine-rich domains of the *su(w^a)* and *tra* RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell* **67**, 335-342.

Loewinger, L. and McKeon, F. (1988). Mutations in the nuclear lamin proteins resulting in their aberrant assembly in the cytoplasm. *EMBO J.* **7**, 2301-2309.

Maeda, Y., Hisatake, K., Kondo, T., Hanada, K., Song, C.-Z., Nishimura, T. and Muramatsu, M. (1992). Mouse rRNA gene transcription factor mUBF requires both HMB-box1 and an acidic tail for

nucleolar accumulation: molecular analysis of the nucleolar targeting mechanism. *EMBO J.* **11**, 3695-3704.

Maridor, G. and Nigg, E. A. (1990). cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar proteins. *Nucl. Acids Res.* **18**, 1286.

Messmer, B. and Dreyer, C. (1993). Requirements for nuclear translocation and nucleolar accumulation of nucleolin of Xenopus laevis. *Eur. J. Cell Biol.* (in press).

Milarski, K. L. and Morimoto, R. I. (1989). Mutational analysis of the human HSP70 protein: distinct domains for nucleolar localization and adenosine triphosphate binding. *J. Cell Biol.* **109**, 1947-1962.

Munro, S. and Pelham, H. R. B. (1984). Use of peptide tagging to detect proteins expressed from cloned genes: deletion mapping functional domains of Drosophila hsp70. *EMBO J.* **3**, 3087-3093.

Munro, S. and Pelham, H. R. B. (1987). A C-terminal signal prevents secretion of luminal ER proteins. *Cell* **48**, 899-907.

Nakagawa, J., Kitten, G. T. and Nigg, E. A. (1989). A somatic cell-derived system for studying both early and late mitotic events in vitro. *J. Cell Sci.* **94**, 449-462.

Newmeyer, D. D. and Forbes, D. J. (1988). Nuclear import can be separated into distinct steps in vitro: nuclear pore binding and translocation. *Cell* **52**, 641-653.

Olson, M. O. J., Orrick, L. R., Jones, C. and Busch, H. (1974). Phosphorylation of acid-soluble nucleolar proteins of novikoff hepatoma ascites cells in vivo. *J. Biol. Chem.* **249**, 2823-2827.

Olson, M. O. and Thompson, B. A. (1983). Distribution of proteins among chromatin components of nucleoli. *Biochemistry* **22**, 3187-3193.

Peculis, B. A. and Gall, J. G. (1992). Localization of the nucleolar protein NO38 in amphibian oocytes. *J. Cell Biol.* **116**, 1-14.

Peter, M., Nakagawa, J., Dorée, M., Labbé, J.-C. and Nigg, E. A. (1990). Identification of major nucleolar proteins as candidate mitotic substrates of cdc2 kinase. *Cell* **60**, 791-801.

Richardson, W. D., Mills, A. D., Dilworth, S. M., Laskey, R. A. and Dingwall, C. (1988). Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. *Cell* **52**, 655-664.

Ris, H. (1989). Three-dimensional imaging of cell ultrastructure with high resolution low-voltage SEM. *Inst. Phys. Conf. Ser.* **98**, 657-662.

Robbins, J., Dilworth, S. M., Laskey, R. A. and Dingwall, C. (1991). Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**, 615-623.

Scheer, U. and Benavente, R. (1990). Functional and dynamic aspects of the mammalian nucleolus. *BioEssays* **12**, 14-21.

Silver, P. A. (1991). How proteins enter the nucleus. *Cell* **64**, 489-497.

Siomi, H., Shida, H., Nam, S. H., Nosaka, T., Maki, M. and Hatanaka, M. (1988). Sequence requirements for nucleolar localization of human T cell leukemia virus type I pX protein, which regulates viral RNA processing. *Cell* **55**, 197-209.

Srivastava, M., Fleming, P. J., Pollard, H. B. and Burns, A. L. (1989). Cloning and sequencing of the human nucleolin cDNA. *FEBS Lett.* **250**, 99-105.

Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A. and Crothers, D. M. (1990). Fragments of the HIV-1 tat protein specifically bind TAR RNA. *Science* **249**, 1281-1285.

Yamasaki, L. and Lanford, R. E. (1992). Nuclear transport: a guide to import receptors. *Trends Cell Biol.* **2**, 123-127.

(Received 4 March 1993 - Accepted 21 April 1993)